AD NO. ___________________________
DTC PROJECT NO. 8-CO-160-UXO-021
REPORT NO. ATC-9775

STANDARDIZED

UXO TECHNOLOGY DEMONSTRATION SITE

MOGULS SCORING RECORD NO. 907

SITE LOCATION:
U.S. ARMY YUMA PROVING GROUND

DEMONSTRATOR:
SKY RESEARCH, INC.
445 DEAD INDIAN MEMORIAL ROAD
ASHLAND, OR 97520

TECHNOLOGY TYPE/PLATFORM:
EM61 MKII/PUSHCART

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

AUGUST 2008

Prepared for:
U.S. ARMY ENVIRONMENTAL COMMAND
ABERDEEN PROVING GROUND, MD 21010-5401

U.S. ARMY DEVELOPMENTAL TEST COMMAND
ABERDEEN PROVING GROUND, MD 21005-5055

DISTRIBUTION UNLIMITED, AUGUST 2008.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
AUG 2008	**Final**	**-**

4. TITLE AND SUBTITLE	
Standardized UXO Technology Domonstration Site Moguls Scoring Record No. 907 (Sky Research, Inc.)	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army Aberdeen Test Center ATTN: TEDT-AT-SLE Aberdeen Proving Ground, MD 21005-5059**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**SAR**	**45**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

DISPOSITION INSTRUCTIONS

TEDT-AT-SLE

MEMORANDUM FOR RECORD

SUBJECT: Operations Security (OPSEC) Review of Paper/Presentation

1. The attached document entitled "Scoring Record No. 907" dated August 2008 is provided for review for public disclosure in accordance with AR 530-1 as supplemented. The document is proposed for public release via the internet.

2. I, the undersigned, am aware of the intelligence interest in open source publications and in the subject matter of the information I have reviewed for intelligence purposes. I certify that I have sufficient technical expertise in the subject matter of this document and that, to the best of my knowledge, the net benefit of this public release outweighs the potential damage to the essential secrecy of all related ATC, DTC, ATEC, Army or other DOD programs of which I am aware.

NAME (Printed)	SIGNATURE	DATE
J. Stephen McClung	J. S. McClung Jr.	August 2008

CONCURRENCE:	NAME (Printed)	SIGNATURE	DATE
Program Mgr/Customer (If not ATC owned technology)	Patrick McDonnell	Patrick McDonnell	27AUG08
Directorate Director	Charles Valz	C Valz	9/8/08
Directorate OPSEC QC and Team Leader	William Burch	William Burch	12AUG08
ATC OPSEC Officer/ Security Manager	Jenell Bigham	Jenell Bigham	18 Sept 2008
Public Affairs Specialist	Crystal Maynard	C Maynard	9/22/08
Technical Director, ATC	John R. Wallace	John R Wallace	22 Sep 08

(Return to ATC PAO for further processing)

DTC GO/SES	N/A		

Encl
as

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
August 2008	Final	21-24, 27 March and 4 April 2006

4. TITLE AND SUBTITLE
STANDARDIZED UXO TECHNOLOGY DEMONSTRATION SITE MOGULS SCORING RECORD NO. 907 (SKY RESEARCH, INC.)

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
McClung, J. Stephen

5d. PROJECT NUMBER
8-CO-160-UXO-021

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Aberdeen Test Center
ATTN: TEDT-AT-SLE
Aberdeen Proving Ground, MD 21005-5059

8. PERFORMING ORGANIZATION REPORT NUMBER
ATC-9775

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Environmental Command
ATTN: IMAE-RTA
Aberdeen Proving Ground, MD 21005-5401

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Same as Item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution unlimited.

13. SUPPLEMENTARY NOTES
None

14. ABSTRACT

This scoring record documents the efforts of Sky Research, Inc. to detect and discriminate inert unexploded ordnance (UXO) utilizing the YPG Standardized UXO Technology Demonstration Site Mogul. This Scoring Record was coordinated by J. Stephen McClung and the Standardized UXO Technology Demonstration Site Scoring Committee. Organizations on the committee include the U.S. Army Corps of Engineers, the Environmental Security Technology Certification Program, the Strategic Environmental Research and Development Program, the Institute for Defense Analysis, the U.S. Army Environmental Command, and the U.S. Army Aberdeen Test Center.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
Unclassified	Unclassified	Unclassified	SAR		19b. TELEPHONE NUMBER *(Include area code)*

ACKNOWLEDGMENTS

Authors:

Rick Fling
Aberdeen Test Support Services (ATSS)
Sverdrup Technology, Inc.
U.S. Army Aberdeen Proving Ground (APG)

Christina McClung
Aberdeen Data Services Team (ADST)
Logistics Engineering and Information Technology Company (Log.Sec)
U.S. Army Aberdeen Proving Ground

Contributors:

William Burch
J. Stephen McClung
Military Environmental Technology Demonstration Center (METDC)
U.S. Army Aberdeen Test Center (ATC)
U.S. Army Aberdeen Proving Ground

Leonardo Lombardo
Aberdeen Test and Support Services
Sverdrup Technology, Inc.
U.S. Army Aberdeen Proving Ground

Patrick McDonnell
Booz Allen Hamilton (BAH)
U.S. Army Environmental Command (USAEC)
U.S. Army Aberdeen Proving Ground

TABLE OF CONTENTS

PAGE

SECTION 4. TECHNICAL PERFORMANCE RESULTS

PAGE

SECTION 5. ON-SITE LABOR COSTS

SECTION 6. COMPARISON OF RESULTS TO OPEN FIELD DEMONSTRATION

SECTION 7. APPENDIXES

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of unexploded ordnance (UXO) require testing so that their performance can be characterized. To that end, Standardized Test Sites have been developed at Aberdeen Proving Ground (APG), Maryland and U.S. Army Yuma Proving Ground (YPG), Arizona. These test sites provide a diversity of geology, climate, terrain, and weather as well as diversity in ordnance and clutter. Testing at these sites is independently administered and analyzed by the government for the purposes of characterizing technologies, tracking performance with system development, comparing performance of different systems, and comparing performance in different environments.

The Standardized UXO Technology Demonstration Site Program is a multi-agency program spearheaded by the U.S. Army Environmental Command (USAEC). The U.S. Army Aberdeen Test Center (ATC) and the U.S. Army Corps of Engineers Engineering Research and Development Center (ERDC) provide programmatic support. The program is being funded and supported by the Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP) and the Army Environmental Quality Technology Program (EQT).

1.2 SCORING OBJECTIVES

The objective in the Standardized UXO Technology Demonstration Site Program is to evaluate the detection and discrimination capabilities of a given technology under various field and soil conditions. Inert munitions and clutter items are positioned in various orientations and depths in the ground.

The evaluation objectives are as follows:

a. To determine detection and discrimination effectiveness under realistic scenarios that vary targets, geology, clutter, topography, and vegetation.

b. To determine cost, time, and manpower requirements to operate the technology.

c. To determine demonstrator's ability to analyze survey data in a timely manner and provide prioritized "Target Lists" with associated confidence levels.

d. To provide independent site management to enable the collection of high quality, ground-truth, geo-referenced data for post-demonstration analysis.

1.2.1 Scoring Methodology

a. The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver-operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}), and those that do not correspond to any known item, termed background alarms.

b. The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the blind grid RESPONSE STAGE, the demonstrator provides the scoring committee with a target response from each and every grid square along with a noise level below which target responses are deemed insufficient to warrant further investigation. This list is generated with minimal processing and, since a value is provided for every grid square, will include signals both above and below the system noise level.

c. The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such and to reject clutter. For the blind grid DISCRIMINATION STAGE, the demonstrator provides the scoring committee with the output of the algorithms applied in the discrimination-stage processing for each grid square. The values in this list are prioritized based on the demonstrator's determination that a grid square is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For digital signal processing, priority ranking is based on algorithm output. For other discrimination approaches, priority ranking is based on human (subjective) judgment. The demonstrator also specifies the threshold in the prioritized ranking that provides optimum performance, (i.e. that is expected to retain all detected ordnance and rejects the maximum amount of clutter).

d. The demonstrator is also scored on EFFICIENCY and REJECTION RATIO, which measures the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. EFFICIENCY measures the fraction of detected ordnance retained after discrimination, while the REJECTION RATIO measures the fraction of false alarms rejected. Both measures are defined relative to performance at the demonstrator-supplied level below which all responses are considered noise, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

e. Based on configuration of the ground truth at the standardized sites and the defined scoring methodology, there exists the possibility of having anomalies within overlapping halos and/or multiple anomalies within halos. In these cases, the following scoring logic is implemented:

(1) In situations where multiple anomalies exist within a single R_{halo}, the anomaly with the strongest response or highest ranking will be assigned to that particular ground truth item.

(2) For overlapping R_{halo} situations, ordnance has precedence over clutter. The anomaly with the strongest response or highest ranking that is closest to the center of a particular ground truth item gets assigned to that item. Remaining anomalies are retained until all matching is complete.

(3) Anomalies located within any R_{halo} that do not get associated with a particular ground truth item are thrown out and are not considered in the analysis.

f. All scoring factors are generated utilizing the Standardized UXO Probability and Plot Program, version 3.1.1.

1.2.2 Scoring Factors

Factors to be measured and evaluated as part of this demonstration include:

a. Response Stage ROC curves:

(1) Probability of Detection (P_d^{res}).

(2) Probability of False Positive (P_{fp}^{res}).

(3) Background Alarm Rate (BAR^{res}) or Probability of Background Alarm (P_{BA}^{res}).

b. Discrimination Stage ROC curves:

(1) Probability of Detection (P_d^{disc}).

(2) Probability of False Positive (P_{fp}^{disc}).

(3) Background Alarm Rate (BAR^{disc}) or Probability of Background Alarm (P_{BA}^{disc}).

c. Metrics:

(1) Efficiency (E).

(2) False Positive Rejection Rate (R_{fp}).

(3) Background Alarm Rejection Rate (R_{BA}).

d. Other:

(1) Probability of Detection by Size and Depth.

(2) Classification by type (i.e., 20-, 40-, 105-mm, etc.).

(3) Location accuracy.

(4) Equipment setup, calibration time and corresponding man-hour requirements.

(5) Survey time and corresponding man-hour requirements.

(6) Reacquisition/resurvey time and man-hour requirements (if any).

(7) Downtime due to system malfunctions and maintenance requirements.

1.3 STANDARD AND NONSTANDARD INERT ORDNANCE TARGETS

The standard and nonstandard ordnance items emplaced in the test areas are listed in Table 1. Standardized targets are members of a set of specific ordnance items that have identical properties to all other items in the set (caliber, configuration, size, weight, aspect ratio, material, filler, magnetic remanence, and nomenclature). Nonstandard targets are inert ordnance items having properties that differ from those in the set of standardized targets.

TABLE 1. INERT ORDNANCE TARGETS

Standard Type	Nonstandard (NS)
20-mm Projectile M55	20-mm Projectile M55
	20-mm Projectile M97
40-mm Grenades M385	40-mm Grenades M385
40-mm Projectile MKII Bodies	40-mm Projectile M813
BDU-28 Submunition	
BLU-26 Submunition	
M42 Submunition	
57-mm Projectile APC M86	
60-mm Mortar M49A3	60-mm Mortar (JPG)
	60-mm Mortar M49
2.75-inch Rocket M230	2.75-inch Rocket M230
	2.75-inch Rocket XM229
MK 118 ROCKEYE	
81-mm Mortar M374	81-mm Mortar (JPG)
	81-mm Mortar M374
105-mm Heat Rounds M456	
105-mm Projectile M60	105-mm Projectile M60
155-mm Projectile M483A1	155-mm Projectile M483A
	500-lb Bomb

HEAT = high-explosive antitank.
JPG = Jefferson Proving Ground.

SECTION 2. DEMONSTRATION

2.1 DEMONSTRATOR INFORMATION

2.1.1 Demonstrator Point of Contact (POC) and Address

POC: Ms. Stacey Kingsbury
540-961-9132

Address: Sky Research, Inc.
445 Dead Indian Memorial Road
Ashland, OR 97520

2.1.2 System Description (provided by demonstrator)

Sky Research is conducting three surveys each at APG and YPG to demonstrate the capabilities of electromagnetic induction (EMI) and magnetometer technologies and our data analysis capabilities. These three surveys include:

a. Survey 1. The active response site and the test sites (calibration lane, blind test grid, and open field scenarios) with Sky Research's EM61-MKII towed array (fig. 1). This survey utilizes an array of five Geonics EM61 MKII sensors deployed with a 0.5-m spacing between each coil. Data are logged using the SKY-DAS at a 10 Hz rate and positioned with the Leica TPS1200 Robotic Total Station (RTS) technology. Additionally, the DAS collects sensor and platform orientation data from the Crossbow AHRS-400 inertial measurement unit (IMU).

b. Survey 2. Active site and test site (calibration lanes, blind test grids, open field, wooded area, moguls, and desert extreme scenarios) with Sky Research's man-portable, quad-sensor magnetometer array; digital compass for orientation; and Leica RTS for positioning. Geometrics G-823 total field cesium vapor magnetometers are being used for this survey. Sky Research deploys this equipment on a low-noise, man-portable, quad-sensor array with an integrated digital compass for sensor orientation information. The G-823 system is configured to stream data at ten samples per channel per second (10 Hz). At a nominal traverse rate of 0.8 m per second (around 3 km/hour), this equates to approximately one sample per 8 cm of forward advance.

c. Survey 3. Calibration lane, blind test grids, and moguls only with Sky Research's gimbaled EM61 MKII, developed via SERDP 1310. The cart is configured to mitigate motion and orientation changes and is positioned with the Leica RTS. This survey deploys the same sensors as survey 1: a Geonics EM61-MKII, Crossbow IMU integrated with the Leica RTS.

Figure 1. Demonstrator's system, EM61 MKII/pushcart.

2.1.3 Data Processing Description (provided by the demonstrator)

a. In addition to standard data processing, we are demonstrating the capability to merge orientation information with sensor data, advanced electromagnetic (EM) and MAG processing capability, and the advanced capability to analyze magnetic and EM data together using the UXOLab software package. This advanced analysis includes the merging of target lists collected by each sensor system and the use of the magnetic data to constrain the EM interpretation via cooperative inversion. Sky Research's standard data processing includes data leveling, statistical data assessment, grid generation, and customized data-filtering to accentuate target signatures. Sky Research uses software from the sensor manufacturers and the UXOLab software developed by the proposed project Principal Investigator, Dr. Stephen Billings, to complete all data-processing tasks.

b. The discrimination methodology we deploy is a variation of the finger-printing method. That is, the response of each anomaly is compared to the response of each item in a library of ordnance items expected to be present in the area. All inversions are performed using the full 3-D position and orientation information of each sensor.

2.1.4 Data Submission Format

Data are submitted for scoring in accordance with data submission protocols outlined in the Standardized UXO Technology Demonstration Site Handbook. These submitted data are not included in this report to protect ground truth information.

2.1.5 Demonstrator Quality Assurance (QA) and Quality Control (QC) (provided by the demonstrator)

QC. The following procedures and logs are used to maximize standardization, repeatability, and control of mapping activities:

a. Equipment Standardization Form: This log documents the daily calibration of each field sensor and navigation system. This form documents the results and analysis of the pre- and post-survey Static Test, Static Spike Test, Cable Shake Test, Backsight, and QC Check Positions.

b. Position Standardization Form: This log documents daily calibration of the real-time kinematic (RTK) Navigation system. Pre-and post-survey results of the 3-Point Navigation Function Test, summary data sampling parameters, and detection of blind seed items are documented.

c. Survey Event Summary Form: This log is used to identify the location of each geophysical survey crew on a daily basis. The log tracks crew members, equipment, filenames, and expected areas to be surveyed. Attached to this daily log are maps of the areas to be surveyed containing the coordinates of benchmarks in the areas as well as the coordinates of each quadrant corner.

d. Data Processing Log: All data from the field is run through a standard data-processing procedure. This procedure is the same for all data and is tracked with the Data Processing Log. This log documents all coordinate transformations, visual data-quality checks, statistical data-quality checks, survey-coverage statistics, interpolation parameters, etc.

e. Target Reanalysis: All targets analyzed as part of the project are subject to review by the project geophysicist. Additionally, a minimum of 10 percent of all targets are reanalyzed by a separate geophysicist to ensure data quality.

QA. QA measures are integrated with the QC activities described above in Section I1 above. Additionally, standardization procedures implemented on a site-specific basis are used to maximize efficiency and to adjust to logistical and schedule requirements. The procedure below is utilized at the site to define the spatial accuracy of the data as well as the repeatability of the sensor readings:

a. A 50-foot-long straight-line transect is established with the positions of the endpoints and midpoint logged via RTS.

b. Wherever possible the traverse line is oriented North to South. Each survey system (sensor and navigation unit) used to collect data is operated over the transect each day following standard procedures as follows:

1) An operator logs background data along the traverse, first heading north from the southern endpoint, and then returning south from the northern endpoint.

2) A metallic pin-flag is placed over the midpoint.

3) The operator logs data along the same path, first traveling north, then returning south.

4) The operator logs data along the same path, first traveling north at a slow pace, then returning south at a significantly more rapid pace.

c. All data lines are downloaded and provided to the site geophysicist for review. These data are examined to determine the repeatability of the pin-flag anomaly amplitude and the repeatability of the positional location of the amplitude peak.

Additionally, for the EM, a static background and spike test is performed twice daily, prior to collecting data and after completion of data collection. This test monitors the instrument background readings, monitors for electronic drift, identifies potential interference, and determines the impulse response and repeatability of measurements over a standard test item. The standard test item is a standard 2-inch-diameter steel trailer hitch ball. For the towed array system, the tow vehicle is turned on during the test. With the instrument held in static position, measurements are recorded for at least 3 minutes. A standard test item is then placed under the center of each coil, and an additional minute of data is recorded. Static background readings for the EM-61 MKII should remain within 2.5 mV of background. Readings for the response of the standard test item should be within 20 percent after subtraction of the sensor baseline response.

For the magnetometer array, a heading calibration and test is performed twice daily, prior to collecting data and after completion of data collection. This test involves traverses across a known point located away from buried UXO or other metallic debris. A 5-m-length of line is walked in eight cardinal directions (N-S, S-N, E-W, W-E, SE-NW, NW-SE, SW-NE, NE-SW). The intersections of each line-direction and each sensor are then compared. If any sensor/line direction combination is found to differ by more than 10 nT, the survey is halted until the reason for this heading-induced error is identified and eliminated.

2.1.6 Additional Records

The following record(s) by this vendor can be accessed via the Internet as Microsoft Word documents at www.uxotestsites.org. The counterpart to this report is the blind grid, scoring record no. 906.

2.2 YPG SITE INFORMATION

2.2.1 Location

YPG is located adjacent to the Colorado River in the Sonoran Desert. The UXO Standardized Test Site is located south of Pole Line Road and east of the Countermine Testing and Training Range. The open field range, calibration grid, blind grid, mogul area, and desert extreme area comprise the 350- by 500-meter general test site area. The open field site is the largest of the test sites and measures approximately 200 by 350 meters. To the east of the open field range are the calibration and blind test grids that measure 30 by 40 meters and 40 by 40 meters, respectively. South of the open field is the 135- by 80-meter Mogul area consisting of a sequence of man-made depressions. The Desert Extreme area is located southeast of the open field site and has dimensions of 50 by 100 meters. The Desert Extreme area, covered with desert-type vegetation, is used to test the performance of different sensor platforms in a more severe desert conditions/environment.

2.2.2 Soil Type

Soil samples were collected at the YPG UXO Standardized Test Site by ERDC to characterize the shallow subsurface (< 3 m). Both surface grab samples and continuous soil borings were acquired. The soils were subjected to several laboratory analyses, including sieve/hydrometer, water content, magnetic susceptibility, dielectric permittivity, X-ray diffraction, and visual description.

Two soil complexes are present within the site: Riverbend-Carrizo and Cristobal-Gunsight. The Riverbend-Carrizo complex is comprised of mixed stream alluvium, whereas the Cristobal-Gunsight complex is derived from fan alluvium. The Cristobal-Gunsight complex covers the majority of the site. Most of the soil samples were classified as either a sandy loam or loamy sand, with most samples containing gravel-size particles. All samples had a measured water content less than 7 percent, except for two that contained 11-percent moisture. The majority of soil samples had water content between 1 to 2 percent. Samples containing more than 3 percent were generally deeper than 1 meter.

An X-ray diffraction analysis on four soil samples indicated a basic mineralogy of quartz, calcite, mica, feldspar, magnetite, and some clay. The presence of magnetite imparted a moderate magnetic susceptibility, with volume susceptibilities generally greater than 100 by 10-5 SI.

For more details concerning the soil properties at the YPG test site, go to www.uxotestsites.org on the web to view the entire soils description report.

2.2.3 Test Areas

A description of the test site areas at YPG is included in Table 2.

TABLE 2. TEST SITE AREAS

Area	Description
Calibration grid	Contains the 15 standard ordnance items buried in six positions at various angles and depths to allow demonstrator equipment calibration.
Blind grid	Contains 400 grid cells in a 0.16-hectare (0.39-acre) site. The center of each grid cell contains ordnance, clutter, or nothing.
Open field	A 4-hectare (10-acre) site containing open areas, dips, ruts, and obstructions, including vegetation.
Mogul	A 2.64 acre area consisting of two areas (the rectangular or driving portion of the course and the triangular section with more difficult, non-drivable terrain). A series of craters (as deep as 0.91m) and trenches (as deep as 0.91m) encompass this section.

SECTION 3. FIELD DATA

3.1 DATE OF FIELD ACTIVITIES (21 through 24, 27 March and 4 April 2006)

3.2 AREAS TESTED/NUMBER OF HOURS

Areas tested and total number of hours operated at each site are summarized in Table 3.

TABLE 3. AREAS TESTED AND NUMBER OF HOURS

Area	Number of Hours
Calibration lanes	6.20
Mogul	24.46

3.3 TEST CONDITIONS

3.3.1 Weather Conditions

A YPG weather station located approximately one mile west of the test site was used to record average temperature and precipitation on a half-hour basis for each day of operation. The temperatures listed in Table 4 represent the average temperature during field operations from 0700 to 1700 hours while precipitation data represents a daily total amount of rainfall. Hourly weather logs used to generate this summary are provided in Appendix B.

TABLE 4. TEMPERATURE/PRECIPITATION DATA SUMMARY

Date, 2006	Average Temperature, ^{o}F	Total Daily Precipitation, in.
21 March	60.26	0.00
22 March	64.69	0.00
23 March	72.59	0.00
24 March	76.69	0.00
27 March	66.67	0.00

3.3.2 Field Conditions

Sky Research surveyed the moguls on 22 through 24 and 27 March 2006. The weather was cool, and the field was dry during the survey.

3.3.3 Soil Moisture

Three soil probes were placed at various locations within the site to capture soil moisture data: blind grid, calibration, desert extreme, open field areas. Measurements were collected in percent moisture and were taken twice daily (morning and afternoon) from five different soil depths (1 to 6 in., 6 to 12 in., 12 to 24 in., 24 to 36 in., and 36 to 48 in.) from each probe. Soil moisture logs are included in Appendix C.

3.4 FIELD ACTIVITIES

3.4.1 Setup/Mobilization

These activities included initial mobilization and daily equipment preparation and breakdown. A three-person crew took 2 hours and 12 minutes to perform the initial setup and mobilization. There was 1 hour and 41 minutes of daily equipment preparation, and end of the day equipment breakdown lasted 47 minutes.

3.4.2 Calibration

Sky Research spent a total of 6 hours and 12 minutes in the calibration lanes, of which 2 hours and 37 minutes were spent collecting data.

3.4.3 Downtime Occasions

Occasions of downtime are grouped into five categories: equipment/data checks or equipment maintenance, equipment failure and repair, weather, demonstration site issues, or breaks/lunch. All downtime is included for the purposes of calculating labor costs (section 5) except for downtime due to demonstration site issues. Demonstration site issues, while noted in the daily log, are considered nonchargeable downtime for the purposes of calculating labor costs and are not discussed. Breaks and lunches are discussed in this section and billed to the total site survey area.

3.4.3.1 Equipment/data checks, maintenance. Equipment data checks and maintenance activities accounted for 7 hours and 13 minutes of site usage time. These activities included changing out batteries and routine data checks to ensure the data was being properly recorded/collected. Sky Research spent an additional 1 hour and 33 minutes for breaks and lunches.

3.4.3.2 Equipment failure or repair. No time was needed to resolve equipment failures that occurred while surveying the mogul.

3.4.3.3 Weather. No weather delays occurred during the survey.

3.4.4 Data Collection

Sky Research spent a total time of 24 hours and 28 minutes in the mogul area, 13 hours and 14 minutes of which was spent collecting data.

3.4.5 Demobilization

The Sky Research survey crew went on to conduct a full demonstration of the site. Therefore, demobilization did not occur until 2 April 2006. On that day, it took the crew 3 hours and 16 minutes to break down and pack up their equipment.

3.5 PROCESSING TIME

Sky Research submitted the raw data from the demonstration activities on the last day of the demonstration, as required. The scoring submittal data were also provided within the required 30-day timeframe.

3.6 DEMONSTRATOR'S FIELD PERSONNEL

Geophysicist: Craig Hyslop
Geophysicist: John Jacobsen
Geophysicist: Rob Mehl

3.7 DEMONSTRATOR'S FIELD SURVEYING METHOD

Sky Research surveyed the moguls in a linear fashion in a north-to-south and a west-to-east direction.

3.8 SUMMARY OF DAILY LOGS

Daily logs capture all field activities during this demonstration and are located in Appendix D. Activities pertinent to this specific demonstration are indicated in highlighted text.

SECTION 4. TECHNICAL PERFORMANCE RESULTS

4.1 ROC CURVES USING ALL ORDNANCE CATEGORIES

The probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive are shown in Figure 2. Both probabilities plotted against their respective background alarm rate are shown in Figure 3. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

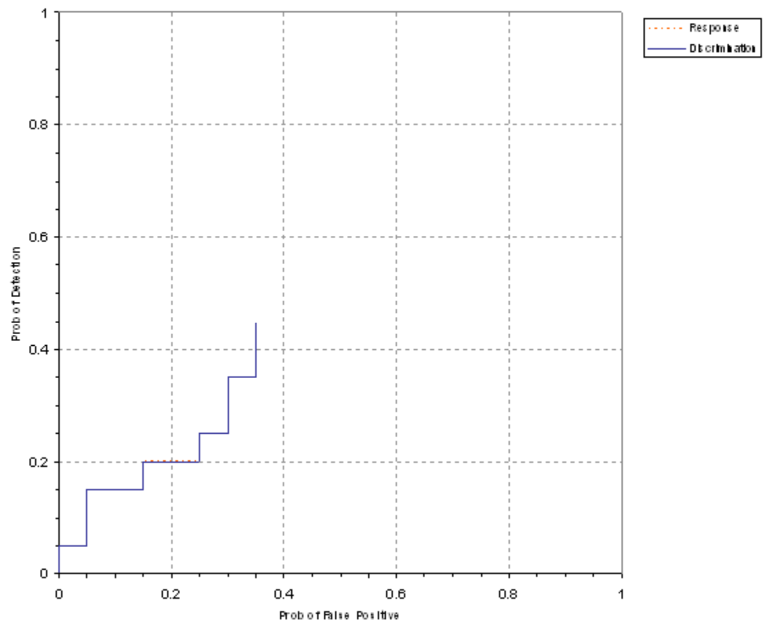

Figure 2. EM61 MKII/pushcart mogul probability of detection for response and discrimination stages versus their respective probability of false positive over all ordnance categories combined.

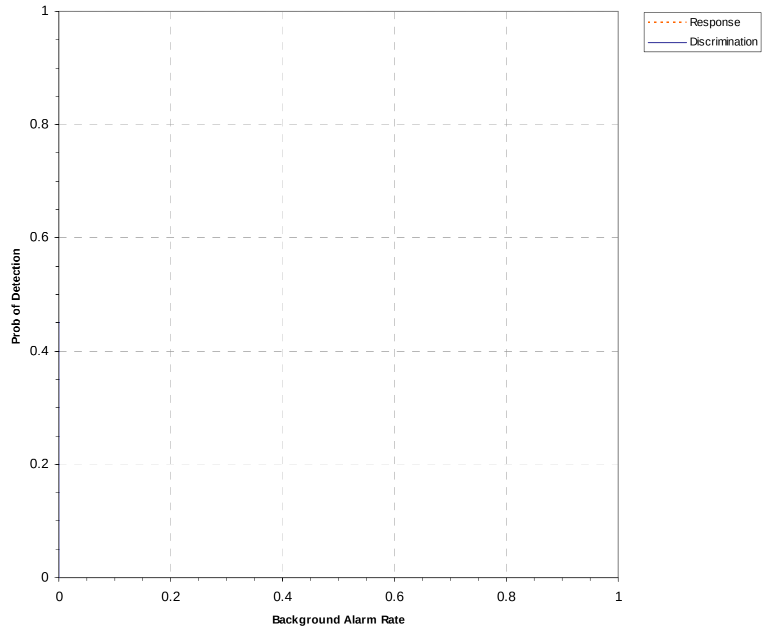

Figure 3. EM61 MKII/pushcart mogul probability of detection for response and discrimination stages versus their respective background alarm rate over all ordnance categories combined.

4.2 ROC CURVES USING ORDNANCE LARGER THAN 20 MM

The probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of false positive when only targets larger than 20 mm are scored are shown in Figure 4. Both probabilities plotted against their respective background alarm rate is shown in Figure 5. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the ground truth.

N/A

Figure 4. EM61 MKII/pushcart mogul probability of detection for response and discrimination stages versus their respective probability of false positive for all ordnance larger than 20 mm.

N/A

Figure 5. EM61 MKII/pushcart mogul probability of detection for response and discrimination stages versus their respective background alarm rate for all ordnance larger than 20 mm.

4.3 PERFORMANCE SUMMARIES

Results for the mogul area test, broken out by size, depth, and nonstandard ordnance are presented in Table 5 (for cost results, see section 5). Results by size and depth include both standard and nonstandard ordnance. The results by size show how well the demonstrator did at detecting/discriminating ordnance of a certain caliber range (see app A for size definitions). The results are relative to the number of ordnance items emplaced. Depth is measured from the geometric center of anomalies.

The RESPONSE STAGE results are derived from the list of anomalies above the demonstrator-provided noise level. The results for the DISCRIMINATION STAGE are derived from the demonstrator's recommended threshold for optimizing UXO field cleanup by minimizing false digs and maximizing ordnance recovery. The lower 90 percent confidence limit on probability of detection and P_{fp} was calculated assuming that the number of detections and false positives are binomially-distributed random variables. All results in Table 5 have been rounded to protect the ground truth. However, lower confidence limits were calculated using actual results.

TABLE 5. SUMMARY OF MOGUL RESULTS FOR EM61 MKII/PUSHCART

Metric	Overall	Standard	Nonstandard	By Size			By Depth, m		
				Small	Medium	Large	< 0.3	0.3 to <1	>= 1
RESPONSE STAGE									
P_d	0.45	0.50	0.40	0.45	0.30	0.75	0.55	0.30	0.30
P_d Low 90% Conf	0.40	0.43	0.30	0.38	0.22	0.58	0.48	0.20	0.08
P_d Upper 90% Conf	0.53	0.59	0.51	0.56	0.44	0.89	0.64	0.42	0.60
P_{fp}	0.35	-	-	-	-	-	0.35	0.40	0.00
P_{fp} Low 90% Conf	0.32	-	-	-	-	-	0.32	0.28	0.00
P_d Upper 90% Conf	0.41	-	-	-	-	-	0.42	0.50	0.68
BAR	0.00	-	-	-	-	-	-	-	-
DISCRIMINATION STAGE									
P_d	0.35	0.35	0.35	0.30	0.25	0.70	0.40	0.20	0.30
P_d Low 90% Conf	0.29	0.28	0.24	0.22	0.18	0.52	0.34	0.13	0.08
P_d Upper 90% Conf	0.41	0.44	0.44	0.38	0.39	0.85	0.49	0.34	0.60
P_{fp}	0.30	-	-	-	-	-	0.30	0.35	0.00
P_{fp} Low 90% Conf	0.28	-	-	-	-	-	0.27	0.26	0.00
P_d Upper 90% Conf	0.37	-	-	-	-	-	0.37	0.47	0.68
BAR	0.00	-	-	-	-	-	-	-	-

Response Stage Noise Level: N/A
Recommended Discrimination Stage Threshold: N/A

Note: The recommended discrimination stage threshold values are provided by the demonstrator.

4.4 EFFICIENCY, REJECTION RATES, AND TYPE CLASSIFICATION

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are reported in Table 6.

TABLE 6. EFFICIENCY AND REJECTION RATES

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	0.75	0.12	0.32
With No Loss of P_d	1.00	0.00	0.09

At the demonstrator's recommended setting, the ordnance items that were detected and correctly discriminated were further scored on whether their correct type could be identified (table 7). Correct type examples include 20-mm projectile, 105-mm HEAT Projectile, and 2.75-inch Rocket. A list of the standard type declaration required for each ordnance item was provided to demonstrators prior to testing. For example, the standard type for the three example items are 20 mmP, 105 H, and 2.75 in. respectively.

TABLE 7. CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS UXO

Size	Percentage Correct
Small	0.0
Medium	0.0
Large	0.0
Overall	0.0

Note: The demonstrator did not attempt to provide type classification.

4.5 LOCATION ACCURACY

The mean location error and standard deviations appear in Table 8. These calculations are based on average missed depth for ordnance correctly identified in the discrimination stage. Depths are measured from the closest point of the ordnance to the surface. For the blind grid, only depth errors are calculated because (X, Y) positions are known to be the centers of each grid square.

TABLE 8. MEAN LOCATION ERROR AND STANDARD DEVIATION (M)

	Mean	Standard Deviation
Northing	0.00	0.19
Easting	0.01	0.19
Depth	0.00	0.00

SECTION 5. ON-SITE LABOR COSTS

A standardized estimate for labor costs associated with this effort was calculated as follows: the first person at the test site was designated supervisor, the second person was designated data analyst, and the third and following personnel were considered field support. Standardized hourly labor rates were charged by title: supervisor at $95.00/hour, data analyst at $57.00/hour, and field support at $28.50/hour.

Government representatives monitored on-site activity. All on-site activities were grouped into one of ten categories: initial setup/mobilization, daily setup/stop, calibration, collecting data, downtime due to break/lunch, downtime due to equipment failure, downtime due to equipment/data checks or maintenance, downtime due to weather, downtime due to demonstration site issue, or demobilization. See Appendix D for the daily activity log. See section 3.4 for a summary of field activities.

The standardized cost estimate associated with the labor needed to perform the field activities is presented in Table 9. Note that calibration time includes time spent in the calibration lanes as well as field calibrations. Site survey time includes daily setup/stop time, collecting data, breaks/lunch, downtime due to equipment/data checks or maintenance, downtime due to failure, and downtime due to weather.

TABLE 9. ON-SITE LABOR COSTS

	No. People	Hourly Wage	Hours	Cost
Initial setup				
Supervisor	1	$95.00	2.20	$209.00
Data analyst	1	57.00	2.20	125.40
Field support	1	28.50	2.20	62.70
Subtotal				**$397.10**
Calibration				
Supervisor	1	$95.00	6.20	$589.00
Data analyst	1	57.00	6.20	353.40
Field support	1	28.50	6.20	176.70
Subtotal				**$1119.10**
Site survey				
Supervisor	1	$95.00	24.46	$2323.70
Data analyst	1	57.00	24.46	1394.22
Field support	1	28.50	24.46	697.11
Subtotal				**$4415.03**

See notes at end of table.

TABLE 9 (CONT'D)

	No. People	Hourly Wage	Hours	Cost
Demobilization				
Supervisor	1	$95.00	3.27	$310.65
Data analyst	1	57.00	3.27	186.39
Field support	1	28.50	3.27	93.20
Subtotal				$590.24
Total				**$6521.47**

Notes: Calibration time includes time spent in the calibration lanes as well as calibration before each data run.

Site Survey time includes daily setup/stop time, collecting data, breaks/lunch, downtime due to system maintenance, failure, and weather.

SECTION 6. COMPARISON OF RESULTS TO OPEN FIELD DEMONSTRATION

No comparison to date. An open field survey was not done during this round of testing.

SECTION 7. APPENDIXES

APPENDIX A. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced ordnance item.

Detection: An anomaly location that is within R_{halo} of an emplaced ordnance item.

Emplaced Ordnance: An ordnance item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., non-ordnance item) buried by the government at a specified location in the test site.

R_{halo}: A pre-determined radius about the periphery of an emplaced item (clutter or ordnance) within which a location identified by the demonstrator as being of interest is considered to be a response from that item. If multiple declarations lie within R_{halo} of any item (clutter or ordnance), the declaration with the highest signal output within the R_{halo} will be utilized. For the purpose of this program, a circular halo 0.5 meters in radius will be placed around the center of the object for all clutter and ordnance items less than 0.6 meters in length. When ordnance items are longer than 0.6 meters, the halo becomes an ellipse where the minor axis remains 1 meter and the major axis is equal to the length of the ordnance plus 1 meter.

Small Ordnance: Caliber of ordnance less than or equal to 40 mm (includes 20-mm projectile, 40-mm projectile, submunitions BLU-26, BLU-63, and M42).

Medium Ordnance: Caliber of ordnance greater than 40 mm and less than or equal to 81 mm (includes 57-mm projectile, 60-mm mortar, 2.75 in. Rocket, MK118 Rockeye, 81-mm mortar).

Large Ordnance: Caliber of ordnance greater than 81 mm (includes 105-mm HEAT, 105-mm projectile, 155-mm projectile, 500-pound bomb).

Shallow: Items buried less than 0.3 meter below ground surface.

Medium: Items buried greater than or equal to 0.3 meter and less than 1 meter below ground surface.

Deep: Items buried greater than or equal to 1 meter below ground surface.

Response Stage Noise Level: The level that represents the point below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the blind grid test area.

Discrimination Stage Threshold: The demonstrator selected threshold level that they believe provides optimum performance of the system by retaining all detectable ordnance and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for n independent trials with the probability p of success and the probability 1-p of failure being the same for each trial. The number of successes x observed in the n trials is an estimate of p and is considered to be a binomially distributed random variable.

RESPONSE AND DISCRIMINATION STAGE DATA

The scoring of the demonstrator's performance is conducted in two stages. These two stages are termed the RESPONSE STAGE and DISCRIMINATION STAGE. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of false positive (P_{fp}) and those that do not correspond to any known item, termed background alarms.

The RESPONSE STAGE scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate ordnance from other anomalies. For the RESPONSE STAGE, the demonstrator provides the scoring committee with the location and signal strength of all anomalies that the demonstrator has deemed sufficient to warrant further investigation and/or processing as potential emplaced ordnance items. This list is generated with minimal processing (e.g., this list will include all signals above the system noise threshold). As such, it represents the most inclusive list of anomalies.

The DISCRIMINATION STAGE evaluates the demonstrator's ability to correctly identify ordnance as such, and to reject clutter. For the same locations as in the RESPONSE STAGE anomaly list, the DISCRIMINATION STAGE list contains the output of the algorithms applied in the discrimination-stage processing. This list is prioritized based on the demonstrator's determination that an anomaly location is likely to contain ordnance. Thus, higher output values are indicative of higher confidence that an ordnance item is present at the specified location. For electronic signal processing, priority ranking is based on algorithm output. For other systems, priority ranking is based on human judgment. The demonstrator also selects the threshold that the demonstrator believes will provide "optimum" system performance, (i.e., that retains all the detected ordnance and rejects the maximum amount of clutter).

Note: The two lists provided by the demonstrator contain identical numbers of potential target locations. They differ only in the priority ranking of the declarations.

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response-stage detections)/(No. of emplaced ordnance in the test site).

Response Stage False Positive (fp^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of False Positive (P_{fp}^{res}): P_{fp}^{res} = (No. of response-stage false positives)/(No. of emplaced clutter items).

Response Stage Background Alarm (ba^{res}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): Blind Grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open Field only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{fp}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can therefore be written as $P_d^{res}(t^{res})$, $P_{fp}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to response-stage data that discriminates ordnance from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to ordnance, as well as those that the demonstrator has high confidence correspond to non-ordnance or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination-stage detections)/(No. of emplaced ordnance in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm (ba^{disc}): An anomaly in a blind grid cell that contains neither emplaced ordnance nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced ordnance or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination-stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination-stage signal strength. These quantities can therefore be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER-OPERATING CHARACERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] Figure A-1 shows how P_d versus P_{fp} and P_d versus BAR are combined into ROC curves. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

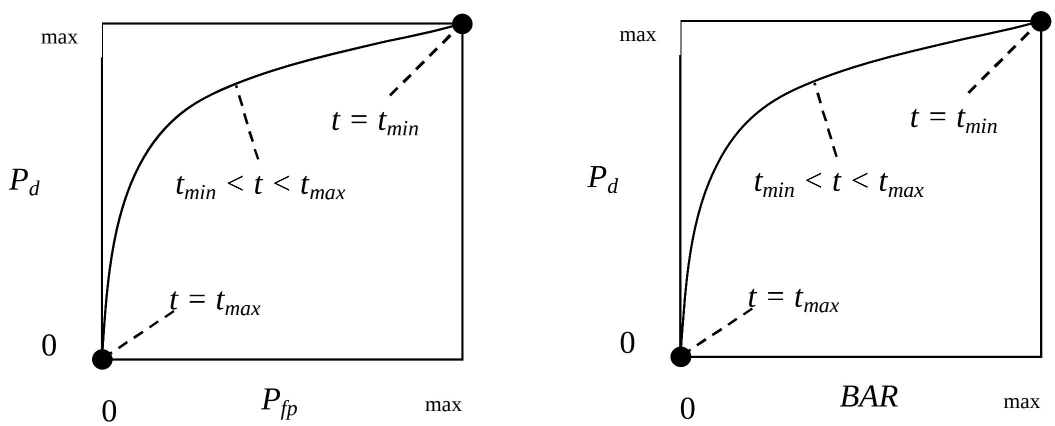

Figure A-1. ROC curves for open field testing. Each curve applies to both the response and discrimination stages.

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a pre-determined and fixed number of detection opportunities (some of the opportunities are located over ordnance and others are located over clutter or blank spots). In an open field scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the open field ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the blind grid test sites are true ROC curves.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of ordnance detections from the anomaly list, while rejecting the maximum number of anomalies arising from non-ordnance items. The efficiency measures the amount of detected ordnance retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum ordnance detectable by the sensor and its accompanying false positive rate or background alarm rate.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$; Measures (at a threshold of interest), the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage tmin) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the ordnance initially detected in the response stage was retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{fp}^{res}(t_{min}^{res})]$; Measures (at a threshold of interest), the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage tmin). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$.
Open field: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})]$).

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON EXPLANATION:

The Chi-square test for differences in probabilities (or 2 x 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations (ref 3).

A 2 x 2 contingency table is used in the Standardized UXO Technology Demonstration Site Program to determine if there is reason to believe that the proportion of ordnance correctly detected/discriminated by demonstrator X’s system is significantly degraded by the more challenging terrain feature introduced. The test statistic of the 2 x 2 contingency table is the

Chi-square distribution with one degree of freedom. Since an association between the more challenging terrain feature and relatively degraded performance is sought, a one-sided test is performed. A significance level of 0.05 is chosen which sets a critical decision limit of 2.71 from the Chi-square distribution with one degree of freedom. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The Chi-square test cannot be used in these instances. Instead, Fischer's test is used and the critical decision limit for one-sided tests is the chosen significance level, which in this case is 0.05. With Fischer's test, if the test statistic is less than the critical value, the proportions are considered to be significantly different.

Standardized UXO Technology Demonstration Site examples, where blind grid results are compared to those from the open field and open field results are compared to those from one of the scenarios, follow. It should be noted that a significant result does not prove a cause and effect relationship exists between the two populations of interest; however, it does serve as a tool to indicate that one data set has experienced a degradation in system performance at a large enough level than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying each of the three progressively more difficult areas using the same system (results indicate the number of ordnance detected divided by the number of ordnance emplaced):

	Blind grid	Open field	Moguls
P_d^{res}	100/100 = 1.0	8/10 = .80	20/33 = .61
P_d^{disc}	80/100 = 0.80	6/10 = .60	8/33 = .24

P_d^{res}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the response stage, all 100 ordnance out of 100 emplaced ordnance items were detected in the blind grid while 8 ordnance out of 10 emplaced were detected in the open field. Fischer's test must be used since a 100 percent success rate occurs in the data. Fischer's test uses the four input values to calculate a test statistic of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open field relative to results from the blind grid using the same system.

P_d^{disc}: BLIND GRID versus OPEN FIELD. Using the example data above to compare probabilities of detection in the discrimination stage, 80 out of 100 emplaced ordnance items were correctly discriminated as ordnance in blind grid testing while 6 ordnance out of 10 emplaced were correctly discriminated as such in open field-testing. Those four values are used to calculate a test statistic of 1.12. Since the test statistic is less than the critical value of 2.71, the two discrimination stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{res}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the response stage, 8 out of 10 and 20 out of 33 are used to calculate a test statistic of 0.56. Since the test statistic is less than the critical value of 2.71, the two response stage detection rates are considered to be not significantly different at the 0.05 level of significance.

P_d^{disc}: OPEN FIELD versus MOGULS. Using the example data above to compare probabilities of detection in the discrimination stage, 6 out of 10 and 8 out of 33 are used to calculate a test statistic of 2.98. Since the test statistic is greater than the critical value of 2.71, the smaller discrimination stage detection rate is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause and effect relationship exists between the change in survey area and degradation in performance, it does indicate that the ability of demonstrator X to correctly discriminate seems to have been degraded by the mogul terrain relative to results from the flat open field using the same system.

APPENDIX B. DAILY WEATHER LOGS

21 March 2006		
Time, 2006	**Temperature, °C**	**Precipitation, in.**
0700	8.2	0.0
0800	11.1	0.0
0900	13.4	0.0
1000	15.1	0.0
1100	15.8	0.0
1200	15.5	0.0
1300	15.1	0.0
1400	18.1	0.0
1500	17.9	0.0
1600	18.4	0.0
1700	18.3	0.0
22 March 2006		
0700	6.9	0.0
0800	11.0	0.0
0900	13.3	0.0
1000	16.4	0.0
1100	17.7	0.0
1200	19.2	0.0
1300	20.3	0.0
1400	21.5	0.0
1500	21.6	0.0
1600	22.0	0.0
1700	22.3	0.0
23 March 2006		
0700	9.6	0.0
0800	13.6	0.0
0900	17.9	0.0
1000	19.9	0.0
1100	21.6	0.0
1200	23.7	0.0
1300	25.5	0.0
1400	26.3	0.0
1500	27.1	0.0
1600	27.5	0.0
1700	27.5	0.0

24 March 2006		
Time, 2006	**Temperature, °C**	**Precipitation, in.**
0700	11.6	0.0
0800	14.8	0.0
0900	18.8	0.0
1000	21.1	0.0
1100	25.5	0.0
1200	26.8	0.0
1300	27.7	0.0
1400	28.6	0.0
1500	29.6	0.0
1600	29.9	0.0
1700	30.0	0.0

APPENDIX C. SOIL MOISTURE

21 March 2006			
Times: 0730 and 1400			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration Area	0 to 6	1.6	1.8
	6 to 12	2.4	2.3
	12 to 24	3.6	3.7
	24 to 36	3.7	3.7
	36 to 48	4.1	4.1
Mogul Area	0 to 6	1.7	1.7
	6 to 12	2.2	2.3
	12 to 24	5.7	5.4
	24 to 36	4.7	4.7
	36 to 48	5.3	5.1
Desert Extreme Area	0 to 6	1.5	1.6
	6 to 12	1.9	1.8
	12 to 24	3.3	3.2
	24 to 36	4.0	4.0
	36 to 48	4.0	4.0
22 March 2006			
Times: 0930 and 1315			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration Area	0 to 6	1.8	1.9
	6 to 12	2.3	2.3
	12 to 24	3.7	3.7
	24 to 36	3.7	3.7
	36 to 48	4.1	4.1
Mogul Area	0 to 6	1.6	1.5
	6 to 12	2.2	2.3
	12 to 24	3.8	3.6
	24 to 36	4.7	4.7
	36 to 48	5.0	4.9
Desert Extreme Area	0 to 6	1.6	1.6
	6 to 12	1.9	1.8
	12 to 24	3.3	3.1
	24 to 36	4.0	4.0
	36 to 48	4.0	4.0

23 March 2006 Times: 0715 and 1300			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration Area	0 to 6	1.7	1.5
	6 to 12	2.4	2.3
	12 to 24	3.6	3.6
	24 to 36	3.7	3.7
	36 to 48	4.1	4.1
Mogul Area	0 to 6	1.6	1.4
	6 to 12	2.6	2.3
	12 to 24	3.8	3.7
	24 to 36	4.7	4.7
	36 to 48	4.9	4.8
Desert Extreme Area	0 to 6	1.5	1.4
	6 to 12	1.8	1.6
	12 to 24	3.3	3.2
	24 to 36	4.0	4.0
	36 to 48	4.0	4.0
24 March 2006 Times: 0745 and 1300			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Calibration Area	0 to 6	1.5	1.8
	6 to 12	2.4	2.3
	12 to 24	3.6	3.8
	24 to 36	3.7	3.7
	36 to 48	4.2	4.1
Mogul Area	0 to 6	1.6	1.5
	6 to 12	2.2	2.3
	12 to 24	3.8	3.6
	24 to 36	4.7	4.7
	36 to 48	5.0	4.7
Desert Extreme Area	0 to 6	3.8	1.6
	6 to 12	3.8	1.9
	12 to 24	3.3	3.2
	24 to 36	4.0	4.0
	36 to 48	4.0	4.0

APPENDIX D. DAILY ACTIVITY LOGS

Date	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration min.	Operation Status	Operational Status - Comments	Track Method	Pattern	Field Conditions
03/21/06	3	CALIBRATION LANES	1148	1400	132	INITIAL SETUP	Assembling the push gimballed EM61-MKII system.	NA	NA	Cloudy, cold
03/21/06	3	CALIBRATION LANES	1400	1418	18	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Ran Clear area to calibrate equipment.	Leica RTS	Linear	Cloudy, cold
03/21/06	3	CALIBRATION LANES	1418	1534	76	COLLECTING DATA	Ran Calibration Grid North to South, West to East using the push gimball; incomplete.	Leica RTS	Linear	Cloudy/windy, cool
03/21/06	3	CALIBRATION LANES	1534	1600	26	DAILY START, STOP	Breakdown end of day	NA	NA	Cloudy/windy, cold
03/22/06	3	CALIBRATION LANES	650	704	14	DAILY START, STOP	Setup of equipment	NA	NA	Clear, cold
03/22/06	3	CALIBRATION LANES	704	840	96	CALIBRATION	Ran Clear area to calibrate equipment.	Leica RTS	Linear	Sunny, cool
03/22/06	3	CALIBRATION LANES	840	1001	81	COLLECTING DATA	Ran Calibration Grid North to South, West to East using the push Gimball; completed.	Leica RTS	Linear	Sunny, warm
03/22/06	3	CALIBRATION LANES	1001	1102	61	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Mapping out the Blind Grid area.	NA	NA	Sunny, warm
03/22/06	3	BLIND TEST GRID	1102	1126	24	COLLECTING DATA	Ran Blind Grid North to South, West to East using the push Gimball; incomplete.	Leica RTS	Linear	Sunny, warm
03/22/06	3	BLIND TEST GRID	1126	1157	31	BREAK/LUNCH	Lunch	NA	NA	Sunny, warm
03/22/06	3	BLIND TEST GRID	1157	1350	113	COLLECTING DATA	Ran Blind Grid North to South, West to East using the push Gimball; completed.	Leica RTS	Linear	Sunny, warm
03/22/06	3	MOGUL	1350	1526	96	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Setting up to run the Mogul grid area.	NA	NA	Sunny, slightly windy, warm

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration min.	Operation Status	Operational Status - Comments	Track Method	Pattern	Field Conditions
03/22/06	3	MOGUL	1610	1624	14	DAILY START, STOP	Breakdown end of day	NA	NA	Sunny, slightly windy, warm
03/23/06	3	MOGUL	615	630	15	DAILY START, STOP	Setup of equipment	NA	NA	Sun coming up, cold
03/23/06	3	MOGUL	630	825	115	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Ran Clear area to calibrate equipment and moving the Leica RTS to run the Mogul Field.	Leica RTS	Linear	Sunny, cool
03/23/06	3	MOGUL	825	1100	155	COLLECTING DATA	Ran Mogul Field North to South, West to East.	Leica RTS	Linear	Sunny, warm
03/23/06	3	MOGUL	1100	1140	40	BREAK/LUNCH	Lunch	NA	NA	Sunny, warm
03/23/06	3	MOGUL	1140	1215	35	COLLECTING DATA	Continued to run Mogul Field North to South, West to East.	Leica RTS	Linear	Sunny, warm
03/23/06	3	MOGUL	1215	1314	59	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Mogul field being cleared of shrubs and brush.	NA	NA	Sunny, warm
03/23/06	3	MOGUL	1314	1557	163	COLLECTING DATA	Continued to run Mogul Field North to South, West to East.	Leica RTS	Linear	Sunny, warm
03/23/06	3	MOGUL	1557	1633	36	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Moving the Leica RTS to run Clear area to validate data.	Leica RTS	Linear	Sunny, warm
03/23/06	3	MOGUL	1633	1650	17	DAILY START, STOP	Breakdown end of day	NA	NA	Sunny, warm
03/24/06	3	MOGUL	620	651	31	DAILY START, STOP	Setup of equipment	NA	NA	Sun coming up, cold
03/24/06	3	MOGUL	651	809	78	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Ran Clear area to calibrate equipment and moving the Leica RTS to run the Mogul Field.	Leica RTS	Linear	Sunny, cool
03/24/06	3	MOGUL	809	1117	188	COLLECTING DATA	Continued to run Mogul Field North to South, West to East.	Leica RTS	Linear	Sunny, warm
03/24/06	3	MOGUL	1117	1210	53	BREAK/LUNCH	Lunch	NA	NA	Sunny, warm

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

Date	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration min.	Operation Status	Operational Status - Comments	Track Method	Pattern	Field Conditions
03/24/06	3	MOGUL	1210	1515	185	COLLECTING DATA	Continued to run Mogul Field North to South, West to East.	Leica RTS	Linear	Sunny, warm
03/24/06	3	MOGUL	1515	1529	14	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Moving the Leica RTS to run Clear area to validate data.	Leica RTS	Linear	Sunny, warm
03/24/06	3	MOGUL	1529	1545	16	DAILY START, STOP	Breakdown end of day	NA	NA	Sunny, warm
03/27/06	3	MOGUL	635	730	55	DAILY START, STOP	Setup of equipment	NA	NA	Overcast, cold
03/27/06	3	MOGUL	730	752	22	DOWNTIME DUE TO EQUIPMENT MAINTENANCE/CHECK	Ran Clear area to calibrate equipment	Leica RTS	Linear	Overcast, cold
03/27/06	3	MOGUL	752	900	68	COLLECTING DATA	Continued to run Mogul Field North to South, West to East; completed.	Leica RTS	Linear	Overcast, cold

Note: Activities pertinent to this specific demonstration are indicated in highlighted text.

APPENDIX E. REFERENCES

1. Standardized UXO Technology Demonstration Site Handbook, DTC Project No. 8-CO-160-000-473, Report No. ATC-8349, March 2002.

2. Aberdeen Proving Ground Soil Survey Report, October 1998.

3. Data Summary, UXO Standardized Test Site: APG Soils Description, May 2002.

4. Yuma Proving Ground Soil Survey Report, May 2003.

5. Practical Nonparametric Statistics, W.J. Conover, John Wiley & Sons, 1980, pages 144 through 151.

APPENDIX F. ABBREVIATIONS

ADST	=	Aberdeen Data Services Team
APG	=	U.S. Army Aberdeen Proving Ground
ATC	=	U.S. Army Aberdeen Test Center
ATSS	=	Aberdeen Test Support Services
BAH	=	Booz Allen Hamilton
EM	=	electromagnetic
EMI	=	electromagnetic induction
ERDC	=	U.S. Army Corps of Engineers Engineering Research and Development Center
ESTCP	=	Environmental Security Technology Certification Program
EQT	=	Army Environmental Quality Technology Program
IMU	=	inertial measurement unit
JPG	=	Jefferson Proving Ground
METDC	=	Military Environmental Technology Demonstration Center
NS	=	nonstandard
POC	=	point of contact
QA	=	quality assurance
QC	=	quality control
ROC	=	receiver-operating characteristic
RTK	=	real-time kinematic
RTS	=	Robotic Total Station
SERDP	=	Strategic Environmental Research and Development Program
USAEC	=	U.S. Army Environmental Center
UXO	=	unexploded ordnance
YPG	=	U.S. Army Yuma Proving Ground

APPENDIX G. DISTRIBUTION LIST

DTC Project No.8-CO-160-UXO-021

Addressee	No. of Copies
Booz Allen Hamilton ATTN: Mr. Patrick McDonnell 4692 Millennium Drive Suite 200 Belcamp, MD 21017-1535	1
Sky Research, Inc. ATTN: Ms. Stacey Kingsbury 445 Dead Indian Memorial Road Ashland, OR 97520	1
Commander U.S. Army Aberdeen Test Center ATTN: TEDT-AT-SLE (Mr. J. Stephen McClung)	1
(Library)	1
400 Colleran Road Aberdeen Proving Ground, MD 21005-5059	
Defense Technical Information Center 8725 John J. Kingman Road, STE 0944 Fort Belvoir, VA 22060-6218	PDF

Secondary distribution is controlled by Commander, U.S. Army Environmental Command, ATTN: IMAE-RTA.

(Page G-2 Blank)